AD NO. ____________________________
DTC PROJECT NO. 8-CO-160-UXO-016
REPORT NO. ATC-9270

SHALLOW WATER

UXO TECHNOLOGY DEMONSTRATION SITE

SCORING RECORD NO. 2

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

DEMONSTRATOR:
TETRA TECH EC, INC.
143 UNION BLVD, SUITE 1010
LAKEWOOD, CO 80228

TECHNOLOGY TYPE/PLATFORM:
DUAL ACOUSTIC AND MAGNETOMETER ARRAY

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

SEPTEMBER 2006

Prepared for:
U.S. ARMY ENVIRONMENTAL CENTER
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, SEPTEMBER 2006.

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

The use of trade names in this document does not constitute an official endorsement or approval of the use of such commercial hardware or software. This document may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
September 2006	Final	24 through 29 October 2005

4. TITLE AND SUBTITLE

STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE SHALLOW WATER SCORING RECORD NO. 2 (TETRA TECH EC, INC.)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rowe, Gary

5d. PROJECT NUMBER

8-CO-160-UXO-016

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Commander
U.S. Army Aberdeen Test Center
ATTN: CSTE-DTC-AT-SL-E
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER

ATC-9270

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Commander
U.S. Army Environmental Center
ATTN: SFIM-AEC-ATT
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution unlimited.

13. SUPPLEMENTARY NOTES

None

14. ABSTRACT

This report documents the efforts of Tetra Tech Ec, Inc. to detect and discriminate inert unexploded ordnance (UXO) using a Dual Acoustic and Magnetometer Array. Testing was conducted at ATC, Standardized Shallow Water UXO Technology Demonstration Site. A description of the tested system and an estimate of survey costs along with the analysis of the system performance are provided.

15. SUBJECT TERMS

Tetra Tech EC, Inc., UXO Standardized Technology Site, Shallow Water, Dual Acoustic and Magnetometer Array, MEC

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Author:

Gary W. Rowe
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Contributors:

William Burch
Military Environmental Technology Demonstration Center
U.S. Army Aberdeen Test Center
U.S. Army Aberdeen Proving Ground

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log. Sec/Tri-S)
U.S. Army Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

SECTION 5. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of munitions and explosives of concern (MEC), i.e., unexploded ordnance (UXO) and discarded military munitions (DMM), require testing so their performance can be characterized. Technologies under development for the detection and discrimination of UXO require independent testing so their performance can be characterized. To that end, the U.S. Army Aberdeen Test Center (ATC) located at Aberdeen Proving Ground (APG), Maryland, has developed a Standardized Shallow Water Test Site. This site provides a controlled environment containing varying water depths, multiple types of ordnance and clutter items, as well as navigational and detection challenges. Testing at this site is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance during system development, and comparing the performance and costs of different systems.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center (USAEC). ATC and the U.S. Army Corps of Engineers Engineering, Research and Development Center (ERDC) provide programmatic support. The Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army Environmental Quality Technology Program (EQT) provided funding and support for this program.

1.2 OBJECTIVE

The objective of the Shallow Water Standardized UXO Technology Demonstration Site is to evaluate the detection and discrimination capabilities of existing and emerging technologies and systems in a shallow water environment. Specifically:

a. To determine the demonstrator’s ability to survey a shallow water area, analyze the survey data, and provide a prioritized “Target List” with associated confidence levels in a timely manner.

b. To determine both the detection and discrimination effectiveness under realistic scenarios that varies ordnance, clutter, and bathymetric conditions.

c. To determine cost, time, and manpower requirements needed to operate the technology.

1.3 CRITERIA

The scoring criteria specified in the Environmental Quality Technology - Operational Requirements Document (EQT-ORD) (app D, ref 1) for: A(1.6.a): UXO Screening, Detection and Discrimination document are presented in Table 1-1. Very little information was available on the capabilities of shallow water detection systems when these criteria were developed. However, they were used in the design of the test site, and the five metrics were used to measure system performance in this report.

TABLE 1-1. SCORING CRITERIA

Metric	Threshold	Objective
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water at a standardized site detected	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water at a standardized site detected
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 10%	Rejection rate of 90% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 0.5%
Reacquisition	Reacquire within 1 meter	Reacquire within 0.5 meter
Cost Rate	$4,000 per acre	$2,000 per acre
Production Rate	5 acres per day	50 acres per day

The ATC shallow water site was designed to evaluate the threshold-detection level of a range of ordnance at the 1-foot + 8-foot requirement. Limited information is available at the objective-detection level. All other measured results will be evaluated against both criteria levels.

1.4 APG SHALLOW WATER SITE INFORMATION

1.4.1 Location

The Aberdeen Area of APG is located in the northeast portion of Maryland on the western shore of the Chesapeake Bay in Harford County. The Shallow Water Test Site is located within a controlled range area of APG.

1.4.2 Soil Type

The area chosen for the shallow water test site was known as Cell No. 3 in a dredge-spoil field. The cell bottom is primarily composed of sediment removed from the Bush River. This is a freshwater site.

1.4.3 Test Areas

a. The test site contains five areas: calibration grid, blind test grid, littoral, open water, and deeper water. Additional detail on each area is presented in Table 1-2. A schematic of the calibration lanes is shown in Figure 1.

TABLE 1-2. TEST AREAS

Area	Description
Calibration Grid	The calibration area contains 15 projectiles, 3 each 40, 60, 81, 105, and 155 mm. One of each projectile type is buried at the projectile diameter to depth ratio shown in Figure 1. This area is designed to provide the user with a sensor library of detection responses for the emplaced targets and an understanding of their resistivity prior to entering the blind test fields. Two “clutter-cloud” target scenarios have been constructed adjacent to this area (fig. 1).
Blind Grid	The blind grid contains 644 detection opportunities. Each grid cell is 2 by 2 m^2. At the center of each cell is either an ordnance item, clutter, or nothing. Surrounding the blind grid on three sides are 3.6-kg (8-lb) shot puts, buried 0.3-meters deep in the sediment. The shot puts can be used as a navigational/ Global Positioning System (GPS) check. The GPS coordinates for the center of each grid and the shot put locations are provided to the vendor prior to testing.
Littoral	This is a sloping area on one side of the pond with vegetation growing into the water line. Water depth ranges from 0.3 to 1.8 meters. It contains a variety of navigational and detection challenges.
Open Water	The open water scenario contains a variety of navigational, detection, and discrimination challenges. Water depth varies from 1.8 to 3.4 meters.
Deeper Water	The water depth in this area varies between 3.4 and 4.3 meters.

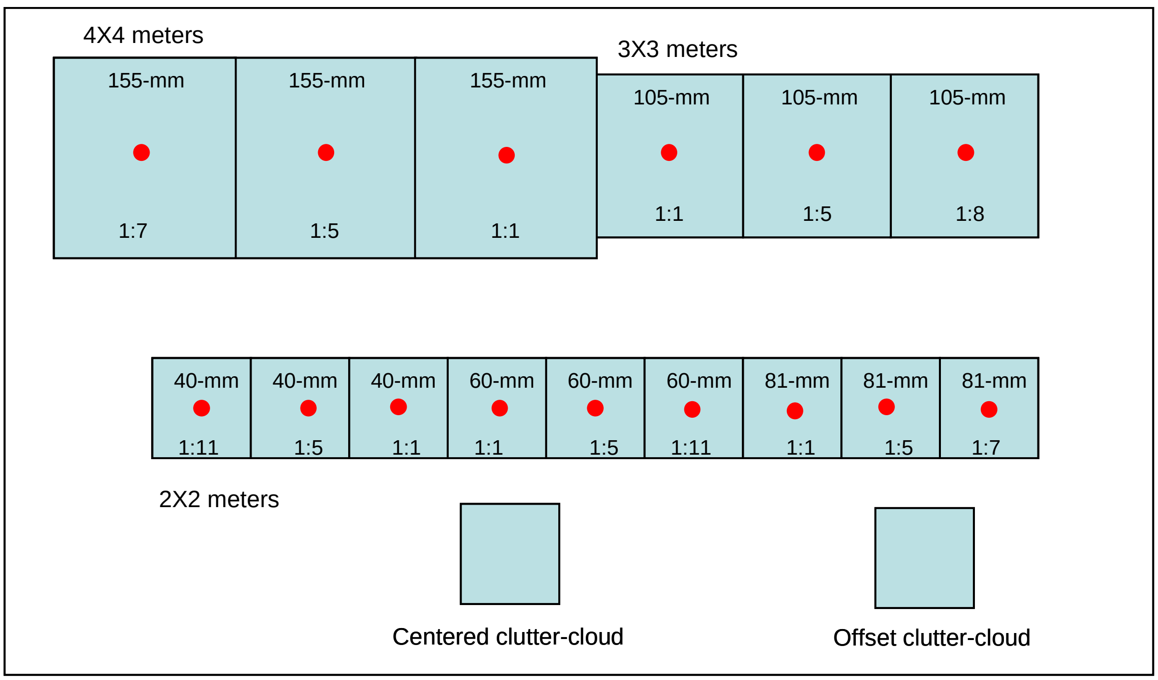

Figure 1. Schematic of the calibration grid.

b. The water depth at this facility during testing is maintained such that the calibration and blind grid areas meet the 2.4-meters (8 ft) detection criterion specified in paragraph 1.3. The test site is approximately 2.8 hectares (6.9 acres) in size.

1.5 GROUND-TRUTH TARGETS

The ground truth is comprised of both inert ordnance and clutter items. The inert ordnance items are listed in Table 1-3. All items were located in storage sites at APG. The items have not been fired or degaussed.

Clutter items fit into one of three categories: ferrous, nonferrous, and mixed-metals. The ferrous and nonferrous items have been further divided into three weight zones as shown in Table 1-4 and distributed throughout all test areas. Most of this clutter is comprised of ordnance components; however, there are also industrial scrap metal and cultural items as well. The mixed-metals clutter is comprised of scrap ordnance items or fragments that have both a ferrous and nonferrous component and could reasonably be encountered in a range area. The mixed-metals clutter was placed in the open water area only.

TABLE 1-3. INERT ORDNANCE TARGETS

Description	**Length, mm**	**Diameter, mm**	**Aspect Ratio, W/L**	**Weight, g**
40-mm L70 Projectile	208	40	0.1923	965
60-mm Mortar M49A2	185	60	0.3243	975
81-mm Mortar M374	528	81	0.1534	3969
81-mm Mortar M821	510	81	0.1588	3338
105-mm Projectile M1	445	105	0.2360	13834
155-mm M107 Projectile	684	155	0.2266	41731
8-inch M104/106	856	203	0.2371	89811

TABLE 1-4. CLUTTER WEIGHT RANGES

Clutter Type	**Weight Range in Grams**		
	Small	**Medium**	**Large**
Ferrous	10 to 510	511 to 2200	> 2201
Nonferrous	10 to 270	275 to 800	> 801

SECTION 2. SYSTEM UNDER TEST

2.1 DEMONSTRATOR INFORMATION

Tetra Tech, as part of their Broad Agency Announcement (BAA) submittal (app D, ref 2), provided the information in paragraphs 2.2 through 2.7 in their detailed test plan. ATC's comments on the demonstrated system are provided in paragraph 2.8.

2.2 SYSTEM DESCRIPTION

A unique system that contains acoustic and magnetic geophysical technology was used. The two acoustic systems were selected in order to be able to identify metal objects on the pond bottom as well as in the shallow pond sediments (through acoustic reflectivity). The acoustic technologies consist of a Sound Metrics Corporation dual frequency imaging and Specialty Devices, Inc., sub-bottom profiler. The proposed magnetic technology was selected to confirm visual indications with the acoustic systems and provide information on the presence of ferrous objects that are potentially out of the detection "window" of the acoustic systems (i.e., buried in the pond sediments). The technology consists of two GEM Systems optically pumped potassium gradiometers (GSMP-40) (fig. 2). Positioning will be accomplished with multiple Leica 1230 Differential Global Positioning System (DGPS) systems, and a Starlink LB-5 swath guidance system was used for system navigation.

The marine vessel that was used for this survey is comprised of two 12-foot fiberglass Jon boats attached together, side by side, with fiberglass supports. There is a space of 2 feet between the boats where the two acoustic technologies are mounted. Several fiberglass poles that extend to a height of approximately 4 to 5 feet were used for GPS antenna mounts. A second fiberglass vessel 8 feet in length supports the magnetic gradiometer system (MGS) array and is pushed 15 to 20 feet ahead of the main vessel. The primary marine vessel is powered by two electric motors. The detection platform is shown in Figure 3.

Two GEM GSMP-40 potassium gradiometers (two sensors per gradiometer) were used in the MGS array. The array geometry consisted of three sensors in a triangular configuration, and one sensor "trailing" the triangular array in a separate horizontal plane. The unique geometric design of the four-sensor MGS allowed the total magnetic field for each sensor to be measured, as well as six magnetic gradients. There were also four analytic signal measurements in two geometric planes that could be automatically calculated from the total field and gradient measurements. The four analytic signal combinations were used to delineate complex magnetic anomalies (e.g., representative of cluttered areas) into their individual constituents so that the anomaly locations are more representative of the individual items present.

Figure 2. GEM GSMP-40 potassium gradiometers array.

Figure 3. Tetra Tech shallow water UXO detection platform.

The increased sensitivity of the MGS over conventional magnetometer systems permits it to be used at or very near the surface of the water, ensuring accurate and precise measurement coordinates using DGPS and efficient data acquisition logistics.

A GEM Systems base station magnetometer was positioned in a magnetically quiet area not more than 0.5 miles from the survey area. The base station magnetometer was used to monitor the geomagnetic field intensity so that these naturally occurring changes could be removed from the data acquired with the MGS array.

The Sound Metrics Corporation High frequency Imaging Sonar (HFIS) (fig. 4) dual frequency imaging sonar operates at 1.1 and 1.8 MHz. For this project, the instrument was used to acquire acoustic image data at 1.8 MHz at a maximum range of 30 meters from the mount location. The x-y-z position of the HFIS unit was determined using a National Marine Electronics Association (NMEA) GPRMC string from a Leica GPS system antenna mounted directly above the HFIS instrument. This permits the image data to be integrated with the Multiple Frequency Sub-Bottom Profiler (MFSBP) and MGS data during analysis.

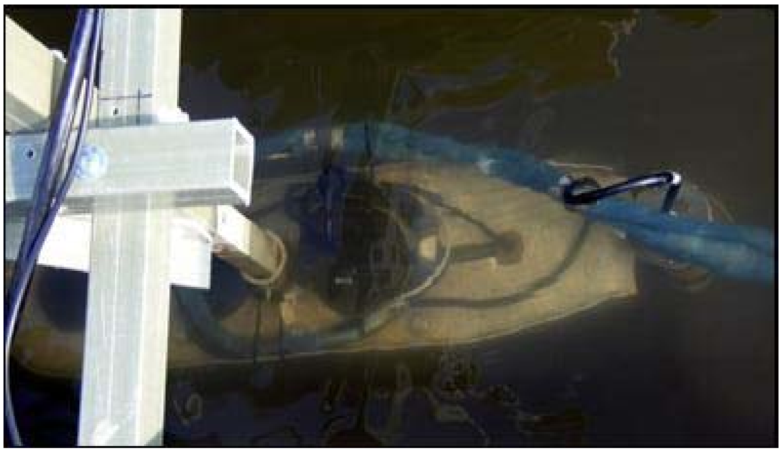

Figure 4. Sonar platform.

The MFSBP system is also deployed on the main acquisition platform at a slight offset (< 1 m) from the Leica GPS unit. When metal is encountered at higher frequencies (e.g., > 24 KHz) the acoustic "reflectivity" of the received signal can be very large compared to the surrounding sediments. The contrast between the metal items of interest and the "background" sediments is maximized when the pond sediments are comprised of soft, lower density grains and particles.

The environmental characteristics of the shallow water detection site (SWDS) permit the use of GPS used in real-time kinematic (RTK) DGPS mode. Due to the unique NMEA code strings required for the acoustic, magnetic, and navigation instrumentation, three to four GPS units and a GPS base station are necessary.

The raw position data for all GPS systems was recorded, and if necessary, post-processed to eliminate potential data gaps.

Navigation was performed using a Starlink LB-5 swath guidance system.

2.3 DEMONSTRATOR POC AND ADDRESS

POC: Mr. Tim Deignan
email: timothy.deignan@tteci.com

Address: Tetra Tech EC, Inc.
143 Union Blvd, Suite 1010
Lakewood, CO 80228

2.4 DEMONSTRATOR'S SITE SURVEYING METHOD

The spatial sample density (line spacing and data recording interval) will be based on the results of the calibration lane data. It was anticipated that a survey line spacing of approximately 3 to 5 feet and the following instrument recording intervals would be sufficient to detect the items of interest for this project:

Recording Intervals
GPS 4 to 5 Hz
MGS 15 to 20 Hz
HFIS 14 to 20 Hz
MFSBP (samples digitized at 100 KHz)

2.5 DEMONSTRATOR QC AND QA

Prior to and at the end of each data acquisition session, a portion of the calibration lane was surveyed, and the data used as a known "benchmark" for the sensitivity and repeatability of the MGS, HFIS, and MFSBP systems. This procedure provided information on the timing systems used in the recording of the magnetic and DGPS equipment so that any latency could be removed from the data during the data processing phase.

At the conclusion of each acquisition session, a portion of the track path from that session was surveyed a second time for QC purposes

2.6 DATA PROCESSING DESCRIPTION

The processed data for each area (i.e., calibration lane, blind grid, open water, littoral, and deep water) was evaluated with respect to anomaly characteristics of each sensor used (e.g., signal intensity, visual identification through HFIS, anomaly size and shape, signal gradients, "noise," and spatial sample density) in order to identify UXO-like items.

Specifically, the HFIS data was used to visually identify the characteristics (length, orientation, overall visual properties, and coordinate position) of items on the pond bottom. The MFSBP data was used to identify the coordinates, distance, and acoustic reflectivity of metal items that existed above and below the pond sediments. The MGS array data provided information on the characteristics of the ferrous metal items present in terms of their distance, ferrous mass, and magnetic dipole direction.

Geosoft Oasis Montaj was used as the data interpretation platform. Color-coded images of the MGS and MFSBP data were generated and compared to the coordinate locations of items visually identified using the HFIS data. The coordinate location of the item was digitized and classified by the interpreter and then output to a dig sheet in the required format.

2.7 DEMONSTRATOR'S SITE PERSONNEL

Project Geophysicist: Mr. Tim Deignan

Data Acquisition Specialists: Adam Maiers
Brian Corbett

2.8 ATC'S SURVEY COMMENTS

Overall, the design of the three-boat system and the instrumentation distribution among the boats was well thought out.

The two electric motors that propelled the three-boat system did not have enough thrust to keep the boat on track in windy conditions. Replacing the two electric motors with a single 3.3-HP gasoline engine solved this problem.

Writable compact disks (CD-R) with the data files from the imaging sonar along with the minimally processed (raw) magnetometer data were provided at the completion of the survey. The final data submission (dig-list) was based on magnetometer data only. Tetra Tech provided the following explanation (app D, ref 3):

"We were unable to fully use all of the data from the five frequency sub-bottom profilers because of software limitations for that particular system (we requested data be available in a certain SEG-Y format for all five frequencies - the vendor is still trying to correctly get this task completed by writing new software using a third party vendor). During data analysis, we used the system depth information and the current software "depthpic"…to view the different frequencies, determine the approximate pond bottom, and correlate this information with the magnetic data. When (and if) the software can convert the multi-frequency data correctly, I expect that we have sufficient x-y data coverage of the area and expanded software analysis tools (as compared to depthpic software) to increase the overall usability of the data.

We did use the data from the DIDSON imaging sonar to select potential items and correlate with the magnetic data; however, it appears that there are extremely few or no metal items that were intended to exist on the pond bottom, however, we do not know this for sure since we do not have the truth data. Basically, the imaging sonar is (and was) implemented to detect items on the pond bottom."

Tetra Tech's pairing of acoustic and metal detecting instrumentation is a unique approach to underwater MEC detection. Problems with acoustic data interpretation, as described above, and detection limitations of the magnetometer array, as described later in this report, show that this particular application of a dual-sensor system was not very successful in terms of underwater MEC detection.

The magnetometer configuration and mounting platform appears better suited for surveying in very shallow water ($\leq$ 1 meter) areas. Detection and discrimination results from the three unbounded test areas show this system performed best in the littoral area.

SECTION 3. SURVEY COST ANALYSIS

3.1 DATES OF SURVEY

The Tetra Tech Inc. Dual Acoustic and Magnetometer Array was tested from 24 through 29 October 2005.

3.2 SITE CONDITIONS

3.2.1 Atmospheric Conditions

An ATC weather station located adjacent to the test site recorded the average temperature and precipitation on an hourly basis for each day of operation. The temperatures listed in Table 3-1 represent the average temperature from 0700 through 1700. The hourly weather logs used to generate this summary are provided in Appendix B.

3.2.2 Water Conditions

Water conditions were monitored using a TIDALITE IV Portable Tide Gauge System©. Data recorded include: water depth and temperature, significant wave height based on the average 1/3 wave height seen over the test period using the Draper/Tucker analysis method, and the full-wave frequency calculated by full-wave mean crossing detection. The values displayed in Table 3-1 were averaged from 0700 through 1700. Detailed information is provided in Appendix B.

TABLE 3-1. SITE CONDITION SUMMARY

Date 05	Air Temperature, °C	Wind, km/h	Water Temperature, °C	Water Depth, m[a]	Significant Wave Height, m	Wave Frequency, Hz
Oct 24	11.5	9.7	12.79	-0.02	0.04	0.14
Oct 25	7.6	9.0	12.00	-0.09	0.10	0.04
Oct 26	10.3	8.0	11.44	0.06	0.18	0.05
Oct 27	9.1	4.6	11.85	-0.25	0.05	0.04
Oct 29	8.6	4.4	11.57	-0.12	0.07	0.04

[a] Variance between the required 2.4 meter test depth and actual test conditions.

3.3 SURVEY ACTIVITIES

The information contained in this section provides an estimate of the time needed and costs associated with surveying an area with this demonstrator's system. This includes data on equipment setup and calibration, site survey and any resurvey time, and downtime due to system malfunctions and maintenance requirements.

3.3.1 Survey Times

a. A government representative monitored and recorded all on-site activities. These activities are grouped into one of eleven categories. The first eight categories are chargeable to the system while the last three are not. Categorizing these activities provides insight into the technical and logistical aspects of the system. The times recorded in each category are then matched with the number of demonstrator personnel, assigned skill levels and a consistent (across vendor) salary to produce an estimate of the survey costs.

(1) Initial setup/mobilization. Starts at the time when the demonstrator's equipment arrives at the survey site and stops when the system is ready to acquire data.

(2) Daily setup/close-up. Monitors time spent mounting and dismounting the equipment each day.

(3) Instrument calibration. Records the amount of time used for daily quality assurance checks, i.e., sensors, GPS data, survey data quality, etc.

(4) Collecting data. Time spent surveying the test area.

(5) Downtime (non-survey time) due to equipment/data checks. Covers time spent trouble shooting equipment or verifying survey tracks.

(6) Downtime (non-survey time) due to equipment failure. Examples are replacing damaged cables, lost communication with base station, or any other failure that prevents surveying. Some weather related failures would fall into this category for example light-emitting diode (LED) displays darken by the sun, wind creating waves to high to survey in, etc.

(7) Downtime (non-survey time) due to maintenance. Battery replacement and memory downloads are typical examples.

(8) Demobilization. Commencement action once the demonstrator has completed the survey and concluded the final on-site check of the test data and ends when the equipment and personnel are ready to leave the site.

(9) Non-chargeable downtime for breaks and lunch. The demonstrator's company policy sets this standard.

(10) Non-chargeable downtime for weather related causes (i.e., lighting, high wet-bulb heat index, and similar events).

(11) Non-chargeable downtime due to ATC range operating requirements. Danger zone conflicts, lack of support personnel, equipment or other ATC caused delays.

b. Appendix C contains the daily log sheets. Table 3-2 summarizes that information to provide insight into the operational, maintenance, and logistical aspects of the system.

TABLE 3-2. TIME ON SITE

Date, 05	Oct 24	Oct 25	Oct 26	Oct 27	Oct 28	Oct 29	Activity Totals, hr
Activity (daily times recorded in minutes)							
Initial setup	435						7.3
Daily setup/close-up	60	120	170	185	110	165	13.5
Instrumentation calibration	55	280	50		65	45	8.3
Data collection		60	180	155	335	310	17.3
Equipment/data checks							0.0
Equipment failure		25	120	115	50	20	5.5
Maintenance							0.0
Demobilization						80	1.3
Breaks and lunch		45					0.8
Weather related							0.0
ATC downtime							0.0
Daily Total, hr	9.2	8.8	8.7	7.6	9.3	10.3	

Note: Task times have been rounded to 5-minute increments.

3.3.2 On-Site Data Collection Costs

The times associated with the 11 activities have been reduced into the three basic components of the evaluation: initial setup, site survey, and pack-up (demobilization). Note that site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather. This combines the actual survey cost with the demonstrator's associated on-site overhead costs.

A standardized estimate for labor costs associated with this effort was then calculated using the following job categories: "supervisor" ($95.00/hr), "data analyst" ($57.00/hr), and "site support" ($28.50/hr). The estimated costs are shown in Table 3-3.

TABLE 3-3. CALCULATED SURVEY COSTS

	No. People	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	7.3	$693.50
Data Analyst	1	$57.00	7.3	$416.10
Site Support	1	$28.50	7.3	$208.05
Subtotal				**$1,317.65**
Site Survey				
Supervisor	1	$95.00	45.4	$4,313.00
Data Analyst	1	$57.00	45.4	$2,587.80
Site Support	1	$28.50	45.4	$1,293.90
Subtotal				**$8,194.70**

TABLE 3-3 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	1.3	$123.50
Data Analyst	1	$57.00	1.3	$74.10
Site Support	1	$28.50	1.3	$37.05
Subtotal				**$234.65**
Total On-site Costs				**$9,747.00**

3.4 COST ANALYSIS

The data collection process described above provides an on-site cost guide to compare the performance of this vendor with any other that has demonstrated at the shallow water site. It is not a true indicator of survey costs. Many other expenses have not been included: travel costs, per Diem, off-site data processing and analysis, company overhead, profit, etc.

Calculating the area surveyed is done by plotting the raw GPS coordinates then combining the sensor swath (line spacing and associated overlap).

To determine the number of acres surveyed per day, the total number of hours spent at the test site (table 3-2) is divided by 8 (converts to 8-hour days). The number of acres is then divided by the number of 8-hour days. The cost per acre is determined by dividing the total survey costs (table 3-3) by the same number of acres. This information is summarized in Table 3-4

TABLE 3-4. SURVEY COSTS

Area Surveyed (Acres[a])	5.57
Time on-site (8-hr days)	6.7
Calculated survey dost (U.S. dollars)	$9,748
Acres per day	0.83
Cost per acre	$1750

[a] Acre = 4047 meters^2

Tetra Tech's survey costs are compared with the EQT-ORD criteria in Table 3-5.

TABLE 3-5. TEST RESULTS - CRITERIA COMPARISION

Metric	Threshold	Objective	TetraTech
Cost Rate	$4,000 per acre	$2,000 per acre	$1,750 per acre
Production Rate	5 acres per day	50 acres per day	0.83 acres per day

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 AREA SURVEYED

4.1.1 Calculated Area

a. Both the test and scoring methodologies require the demonstrator to survey 100 percent of each of the four test areas (blind grid, open water, littoral, and deeper water). Scoring a partially surveyed area alters the ordnance and clutter sample sizes, and test area boundaries and decreases the statistical confidence in the performance statements made for that area. Allowing partial scoring decreases the validity of performance comparisons made between multiple test areas for a single demonstrator and comparisons made between multiple demonstrators for a single test area.

b. Realizing that some systems may not be able to survey 100 percent of a given test area, a ranking system was established. The percent coverage for a given test area is determined by first plotting the raw GPS coordinates combined with the sensor swath (line spacing and associated overlap), calculating the area surveyed, and then comparing that surveyed area to the total test area.

$$\frac{\text{Section Surveyed}}{\text{Test Area Size}} \times 100 = \% \text{ surveyed}$$

c. The demonstrator's system is always scored against the complete ground truth for a given test area regardless of the percentage covered.

4.1.2 Area Assessment

The ranking system and survey results are presented in Table 4-1.

TABLE 4.1. SURVEY RANKING SYSTEM AND RESULTS

Ranking System		**Survey Results**		
% Area Covered	**Ranking**	**Test Area**	**% Area Covered**	**Data Usage**
95 to 100	Met	Blind Grid	100	Direct comparison between systems and areas.
		Open Water	99	
		Deeper Water	98	
90 to 94	Generally met			Comparison between systems and areas. A small negative bias is contained in the reported numbers (bias not quantified in this report).
50 to 89	Partially met	Littoral	88	Reported, not compared between systems or areas. A large negative bias is contained in the reported numbers (bias not quantified in this report).
0 to 49	Not met			Not scored/not reported.

4.2 SYSTEM SCORING PROCEDURES

a. The scoring entities used in this program are predicated on knowing the composition and location of every detectable item in an area. The deeper water area is the one exception. Ground-truth targets were placed in this area without a pre-survey and clearing operation. Therefore, only the system's probability of detection (P_d) is evaluated in this area.

b. The best indicator of survey performance is the blind grid. This area provides a statically valid, controlled environment in which the demonstrator must provide a response (ordnance, clutter, or blank) at each of the 644 locations. Comparison of the response and discrimination lists to the ground truth in this area both determines the range of ordnance the system can reliably detect and establishes the baseline to which system performance in all other test areas is measured.

c. The scoring terms and definitions along with an explanation of the receiver operating characteristics (ROC) curve development and the Chi-square analysis used in this report are in Appendix C.

d. Demonstrator performance is scored in two stages: response and discrimination.

e. The response stage scoring evaluates the ability of the demonstrator's system to detect emplaced ground-truth targets without regard to discriminating ordnance from clutter. In this stage, the GPS locations and signal strengths of all anomalies that the demonstrator has deemed sufficient for further investigation and/or processing are reported. This list is generated with minimal processing, i.e., associating signal strength with GPS location, and includes only signals that are above the system noise level.

f. The discrimination stage evaluates the demonstrator's ability to segregate ordnance from clutter. The same GPS locations reported in the response stage anomaly list are evaluated based on the demonstrator's discrimination process (para 2.1.5). A discrimination stage list is generated and prioritized based on the demonstrator's determination that an anomaly is more likely to be ordnance rather than clutter. Typically, higher output values indicate a higher confidence that an ordnance item is present at a specified location. The demonstrator then specifies the threshold value for the prioritized ranking that provides optimum system performance. This value is the discrimination stage threshold.

g. Both the response and discrimination lists contain an identical number of potential target locations. They differ only in the priority ranking of the declarations.

h. Within both of these stages, the following entities are measured:

(1) P_d.

(2) Probability of false positive (P_{fp}).

(3) Probability of background alarm (P_{ba})/background alarm rate (BAR).

4.2.1 <u>ROC curves</u>

Based on the entire range of ground-truth targets used at this site, ROC curves were generated for both the response and discrimination stages. In both stages, the probability of detection versus false alarm rates is plotted. False alarms are divided into two groups; those anomalies that correspond to emplaced clutter items, thereby measuring the P_{fp}, and anomalies that do not correspond to any known item, termed background alarms (P_{ba}) in the blind grid area and BAR in all other areas.

The ROC curves for the response and discrimination stages for all areas surveyed are shown in Figures 5 through 10. Horizontal lines illustrate the system performance at the demonstrator's recommended noise level during the response stage, or discrimination threshold level in the discrimination stage. The point where the curve crosses the horizontal line defines the subset of targets the demonstrator recommends digging.

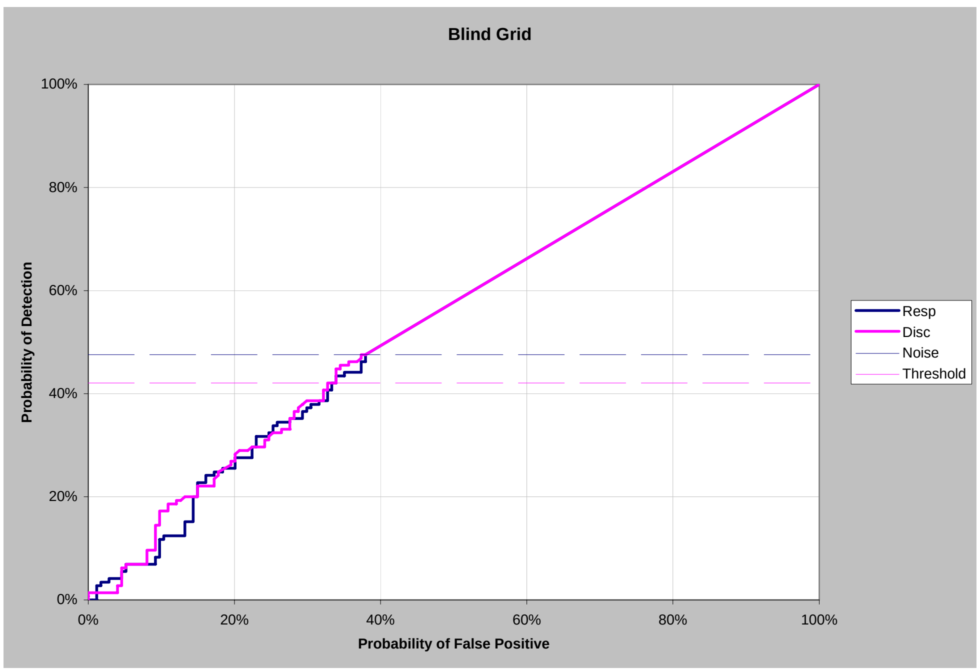

Figure 5. Blind grid P_d versus P_{fp}.

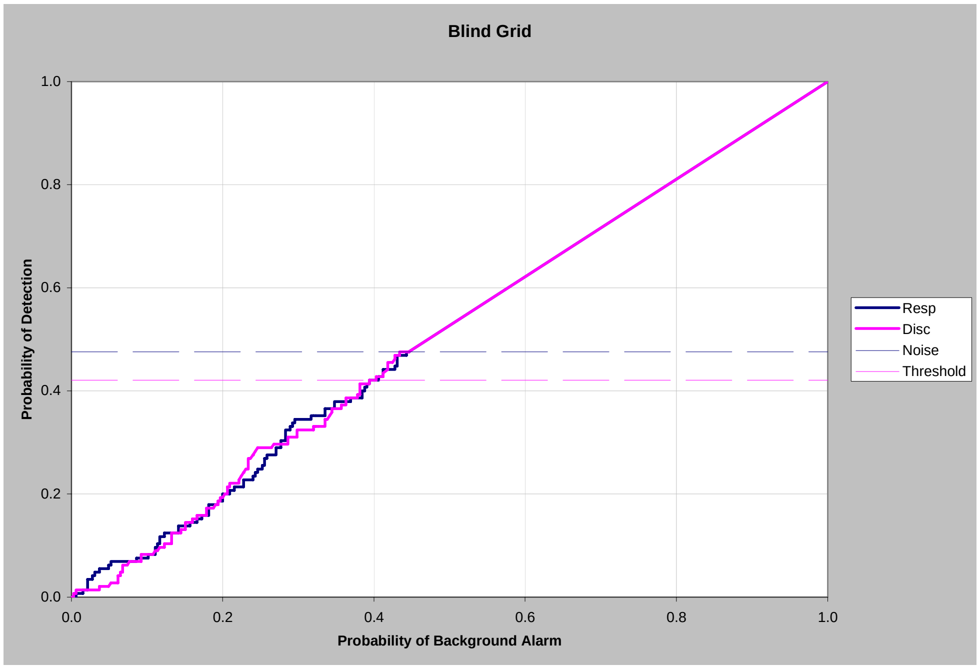

Figure 6. Blind grid P_d versus P_{ba}.

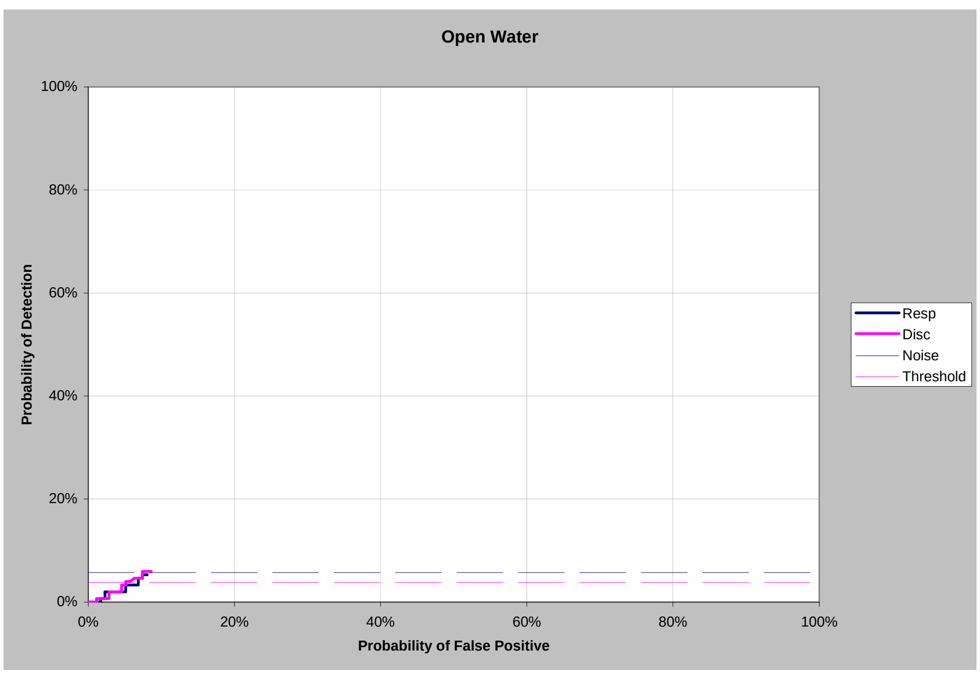

Figure 7. Open water P_d versus P_{fp}.

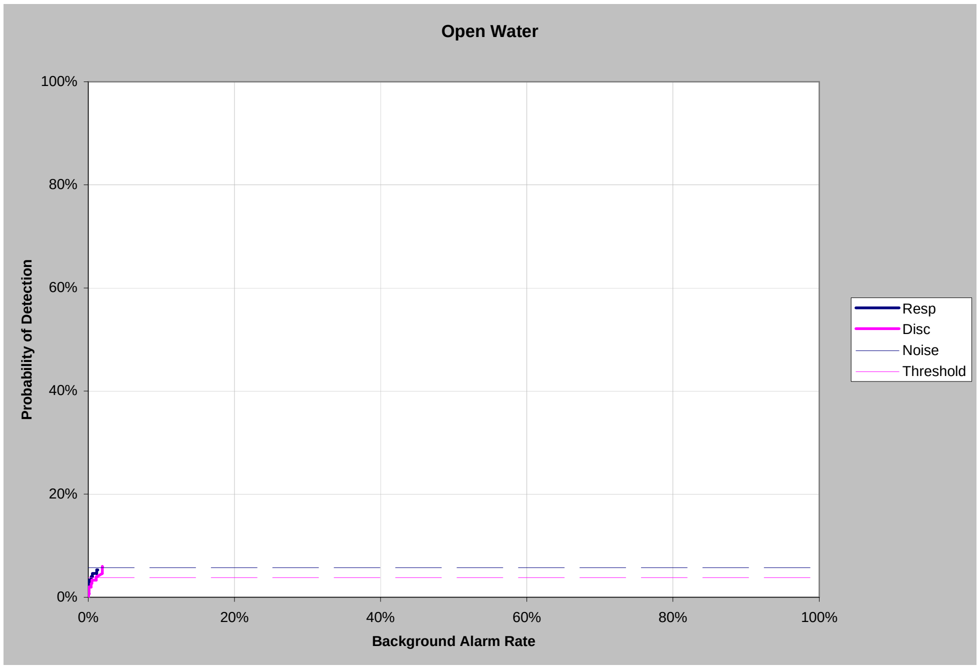

Figure 8. Open water P_d versus BAR.

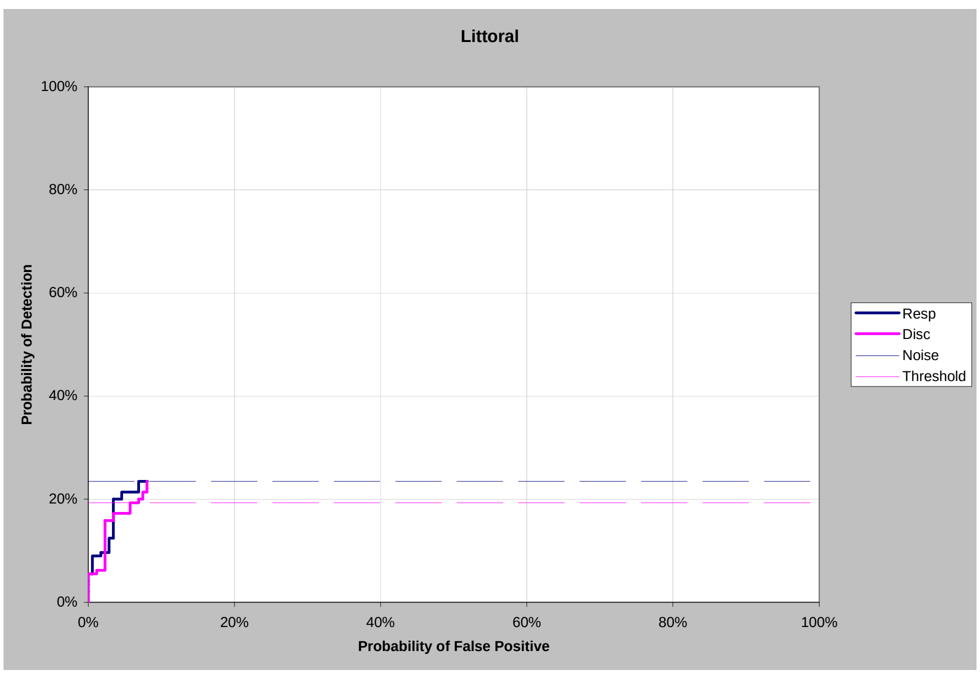

Figure 9. Littoral P_d versus P_{fp}.

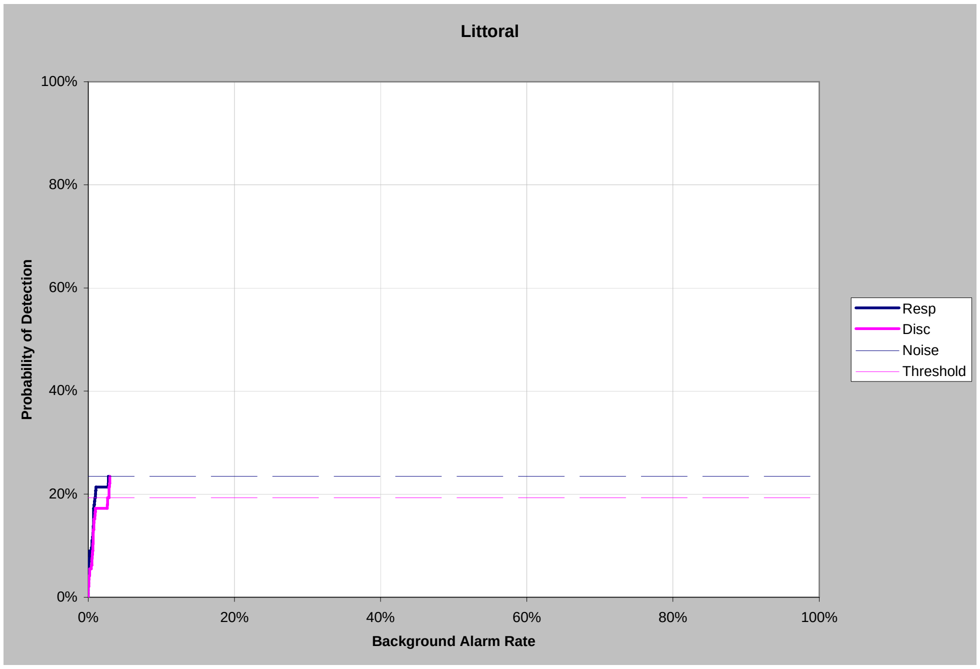

Figure 10. Littoral P_d versus BAR.

4.2.2 Detection Results

Detection results, broken out by stage, area surveyed, and ordnance size are in Table 4-2. The results by size indicate how well the demonstrator did at detecting/discriminating ordnance of a given caliber. Overall results summarize ordnance detection over a given area. All values were calculated assuming the number of detections is a binomially distributed random variable. These results are reported at the 90-percent reliability/95-percent confidence levels unless otherwise noted

TABLE 4-2. SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40 mm	60 mm	81 mm	105 mm	155 mm	8 inch
Blind Grid							
Response Stage							
P_d	47.6%	58.6%	41.4%	51.7%	37.9%	48.3%	
P_d Lower 90% Confidence	42.0%	45.0%	28.8%	38.4%	25.7%	35.1%	
P_{fp}	37.9%						
P_{fp} Lower 90% Confidence	33.1%						
P_{ba}	44.6%						
Discrimination Stage							
P_d	42.1%	44.8%	37.9%	44.8%	37.9%	44.8%	
P_d Lower 90% Confidence	36.6%	31.9%	25.7%	31.9%	25.7%	31.9%	
P_{fp}	32.8%						
P_{fp} Lower 90% Confidence	28.1%						
P_{ba}	39.7%						
Open Water							
Response Stage							
P_d	5.7%	10.3%	6.9%	3.4%	6.9%	2.9%	0.0%
P_d Lower 90% Confidence	3.5%	3.9%	1.8%	0.4%	1.8%	0.3%	0.0%
P_{fp}	7.4%						
P_{fp} Lower 90% Confidence	5.1%						
BAR m^{-2}	0.020						
Discrimination Stage							
P_d	3.8%	3.4%	6.9%	3.4%	3.4%	2.9%	0.0%
P_d Lower 90% Confidence	2.0%	0.4%	1.8%	0.4%	0.4%	0.3%	0.0%
P_{fp}	4.9%						
P_{fp} Lower 90% Confidence	3.1%						
BAR m^{-2}	0.012						
Littoral Region							
Response Stage							
P_d	23.4%	10.3%	3.4%	17.2%	44.8%	41.4%	
P_d Lower 90% Confidence	18.9%	3.9%	0.4%	8.6%	31.9%	28.8%	
P_{fp}	8.0%						
P_{fp} Lower 90% Confidence	5.5%						
BAR m^{-2}	0.030						
Discrimination Stage							
P_d	19.3%	10.3%	3.4%	17.2%	34.5%	31.0%	
P_d Lower 90% Confidence	15.1%	3.9%	0.4%	8.6%	22.6%	19.7%	
P_{fp}	6.9%						
P_{fp} Lower 90% Confidence	4.5%						
BAR m^{-2}	0.028						
Deeper Water							
Response Stage							
P_d	6.9%					6.9%	
P_d Lower 90% Confidence	1.8%					1.8%	
Discrimination Stage							
P_d	6.9%					6.9%	
P_d Lower 90% Confidence	1.8%					1.8%	
Response Stage Noise Level: 0.1 for all areas							
Recommended Discrimination Threshold: Blind Grid =3 Open Water, Littoral and Deeper Water =1							

4.2.3 System Discrimination

Using the demonstrator's recommended setting, the items that were detected and correctly classified as ordnance were further evaluated as to whether the demonstrator could correctly identify the ordnance type. The list of ground-truth ordnance items was provided to the demonstrator prior to testing.

Tetra Tech's "dig-list" discriminated between ordnance and clutter, but not between ordnance types. The latter was an optional requirement.

4.2.4 System Effectiveness

Efficiency and rejection rates were calculated to quantify the discrimination ability at two specific points of interest on the ROC curve: at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and the operator-selected threshold. These values are presented in Table 4-3.

TABLE 4-3. EFFICIENCY AND REJECTION RATES

	Efficiency	**False Positive Rejection Rate**	**Background Alarm Rejection Rate**
Blind Grid			
At operating point	0.88	0.14	0.11
With no loss of P_d	1.00	0.02	0.03
Open Water			
At operating point	0.67	0.33	0.40
With no loss of P_d	1.00	0.13	0.02
Littoral Region			
At operating point	0.82	0.14	0.06
With no loss of P_d	1.00	0.00	0.01

4.2.5 Chi-Square Analysis

Chi-square 2 by 2 Contingency Test for comparison between ratios was used to compare performance across test areas with regard to P_d^{res}, P_d^{disc}, P_{fp}^{res}, and P_{fp}^{disc}, efficiency, and false alarm rejection rate. A one-sided Chi-square significance test at the 0.05 significance level was used to compare the survey results from the blind grid to both the open water and littoral areas. A two-sided test at the 0.10 significance level was used to compare the open water results to those obtained in the littoral zone. The intent of the comparison was to determine if the features introduced in each test site had a degrading effect on the performance of the sensor system. These results are presented in Table 4-4.

TABLE 4-4. CHI-SQUARE SIGNIFICANCE TEST RESULTS

Metric	Overall	By Projectile Caliber				
		40 mm	60 mm	81 mm	105 mm	155 mm
[a]**Blind Grid - Open Water Comparison**						
P_d^{res}	SIG	SIG	SIG	SIG	SIG	SIG
P_d^{disc}	SIG	SIG	SIG	SIG	SIG	SIG
P_{fp}^{res}	SIG					
P_{fp}^{disc}	SIG					
Efficiency	Not					
Rejection rate (R_{fp})	Not					
[a]**Blind Grid - Littoral Region Comparison**						
P_d^{res}	SIG	SIG	SIG	SIG	Not	Not
P_d^{disc}	SIG	SIG	SIG	SIG	Not	Not
P_{fp}^{res}	SIG					
P_{fp}^{disc}	SIG					
Efficiency	Not					
Rejection rate (R_{fp})	Not					
[b]**Open Water - Littoral Comparison**						
P_d^{res}	LR	Not	Not	Not	LR	LR
P_d^{disc}	LR	Not	Not	Not	LR	LR
P_{fp}^{res}	Not					
P_{fp}^{disc}	Not					
Efficiency	Not					
Rejection rate (R_{fp})	Not					
SIG = Significant Not = Not Significant						

[a] One-sided comparison performed at the 0.05 significance level.
[b] Two-sided comparison performed at the 0.10 significance level. The greater (better performance area) of the two areas is noted when a significant difference is detected.

4.2.6 Location Accuracy

The data points in the following scatter-graphs represent the coordinates of ordnance items in the open water and littoral test areas that were first detected in the response stage within a 0.5-meter radius of their true positions, then correctly identified as ordnance in the discrimination stage. The maximum error represents the 0.5-meter detection limit. The mean error represents the statistical mean of the sample considered.

A visual assessment of the littoral graph indicates location error is a randomly distributed as opposed to a systematic error. The open water graph does not contain enough data to make an analysis or statement regarding location accuracy.

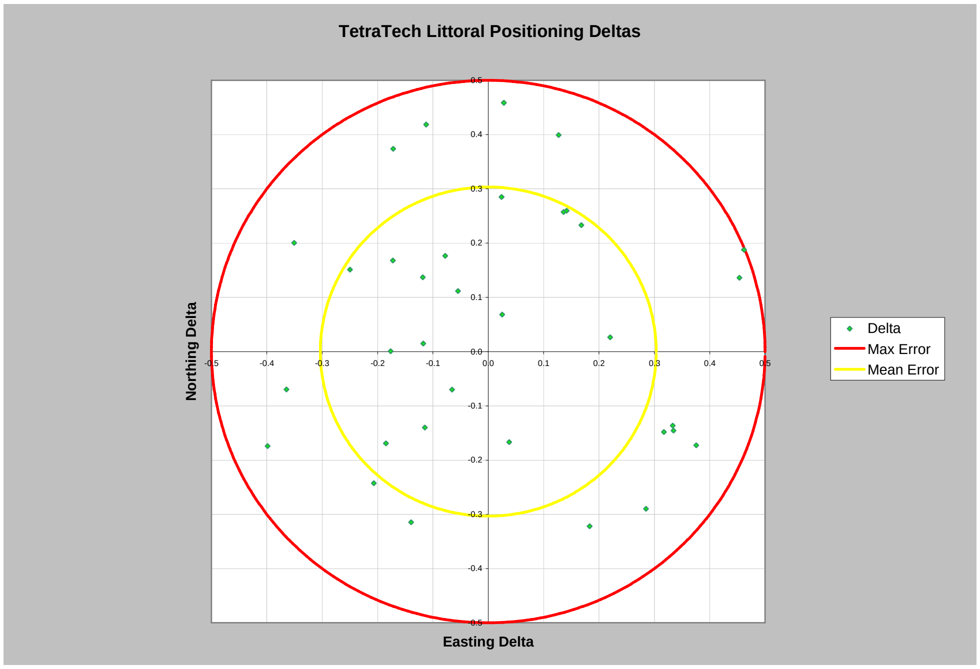

Figure 11. Tetra Tech littoral positional deltas.

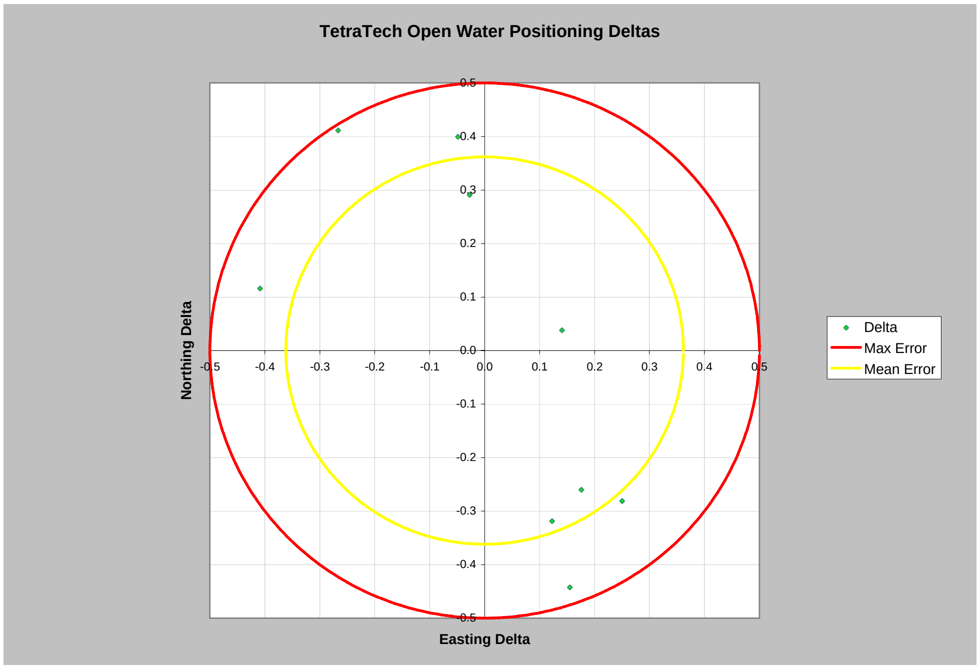

Figure 12. Tetra Tech open water positional deltas.

The comparison between the results obtained during testing and the EQT-ORD criteria are presented in Table 4-5.

TABLE 4-5. TEST RESULTS - CRITERIA COMPARISION

Metric	Threshold	Objective	Tetra Tech by Area	
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water.	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water.	Blind Grid	47.6%
			Open Water	5.7%
			Littoral	23.4%
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter.	Rejection rate of 90% of emplaced non-UXO clutter.	Blind Grid	14%
			Open Water	33%
			Littoral	14%
	Maximum false negative rate of 10%.	Maximum false negative rate of 0.5%.	Not assessed. An analytical procedure is not available to address this criterion.	
Reacquisition	Reacquire within 1 meter.	Reacquire within 0.5 meter.	The reported detection values are based on ordnance items identified within 0.5meters of the georeferenced ground-truth targets.	

Note: The blind grid and open water areas are in general accordance with the threshold requirements.

SECTION 5. APPENDIXES

APPENDIX A. TEST CONDITIONS LOG

ATMOSPHERIC CONDITIONS

Date mm/dd/yy	Time, EDT	Average Wind Direction, deg	Average Wind Speed, km/h	Wind Direction Average Standard Deviation, deg	Peak Wind Speed, km/h	Average Temperture, ^{o}C
10/24/05	0700	71	6.1	19	13.0	8.7
	0800	66	6.1	19	13.8	8.7
	0900	66	6.9	21	18.3	9.4
	1000	66	8.5	25	24.0	11.1
	1100	66	10.0	25	25.1	11.6
	1200	76	11.3	27	26.1	12.8
	1300	74	11.3	25	26.9	12.9
	1400	76	10.0	26	26.1	12.9
	1500	74	12.7	25	31.2	13.1
	1600	59	12.6	22	26.2	12.7
	1700	61	11.7	23	29.1	12.3
10/25/05	0700	4	8.4	17	24.9	7.5
	0800	351	9.4	16	27.4	7.6
	0900	337	9.3	15	26.1	7.6
	1000	342	8.4	20	25.4	7.5
	1100	333	9.1	18	28.5	7.3
	1200	345	11.5	16	35.2	7.4
	1300	344	10.9	17	33.1	7.7
	1400	344	9.3	17	29.6	7.8
	1500	343	8.8	16	25.6	7.7
	1600	335	6.7	16	24.1	7.7
	1700	340	6.9	17	20.6	7.6
10/26/05	0700	319	6.5	18	21.7	6.1
	0800	289	5.6	20	18.3	5.8
	0900	276	6.4	17	17.7	6.4
	1000	293	5.8	21	20.3	8.3
	1100	305	8.0	21	27.8	10.2
	1200	304	8.5	19	31.4	11.3
	1300	301	8.5	20	29.3	12.3
	1400	295	10.1	19	30.1	12.7
	1500	300	9.7	22	30.9	13.2
	1600	299	9.1	20	29.1	13.4
	1700	291	9.3	19	28.8	13.3
10/27/05	0700	298	1.4	24	6.1	2.7
	0800	352	1.0	19	7.4	2.5
	0900	332	2.9	30	14.0	5.1
	1000	330	5.4	19	20.0	8.3
	1100	337	5.6	18	17.4	10.1
	1200	325	4.6	24	14.3	11.3
	1300	314	4.7	27	16.1	12.4
	1400	336	6.2	21	20.0	12.4
	1500	336	7.5	22	21.2	12.1
	1600	326	7.0	17	21.2	11.7
	1700	314	4.2	20	14.8	11.2

(CONT'D)

Date mm/dd/yy	Time, EDT	Average Wind Direction, deg	Average Wind Speed, km/h	Wind Direction Average Standard Deviation, deg	Peak Wind Speed, km/h	Average Temperture, ^{o}C
10/28/05	0700	13	4.3	10	12.4	4.9
	0800	16	3.4	12	10.8	4.8
	0900	57	1.9	38	8.2	4.4
	1000	19	3.0	39	10.8	6.2
	1100	30	4.9	26	16.3	8.8
	1200	67	3.6	49	12.4	10.3
	1300	49	3.4	55	13.8	11.1
	1400	325	4.6	34	16.4	11.1
	1500	328	6.1	18	16.7	11.3
	1600	345	7.3	18	20.3	11.1
	1700	353	5.4	28	22.8	10.9
10/29/05	0700	328	2.6	17	10.1	4.1
	0800	325	3.8	16	13.2	4.6
	0900	341	5.4	19	15.6	5.3
	1000	354	6.1	17	18.8	6.6
	1100	5	6.3	21	20.9	8.4
	1200	350	8.0	19	24.8	9.5
	1300	343	8.2	19	23.5	9.2
	1400	330	7.8	20	28.6	9.6
	1500	315	8.4	20	26.9	9.6
	1600	325	10.1	20	33.3	10.2
	1700	320	8.4	20	29.1	10.4

The water conditions during the Tetra Tech survey are shown in Figures A-1 through A-5.

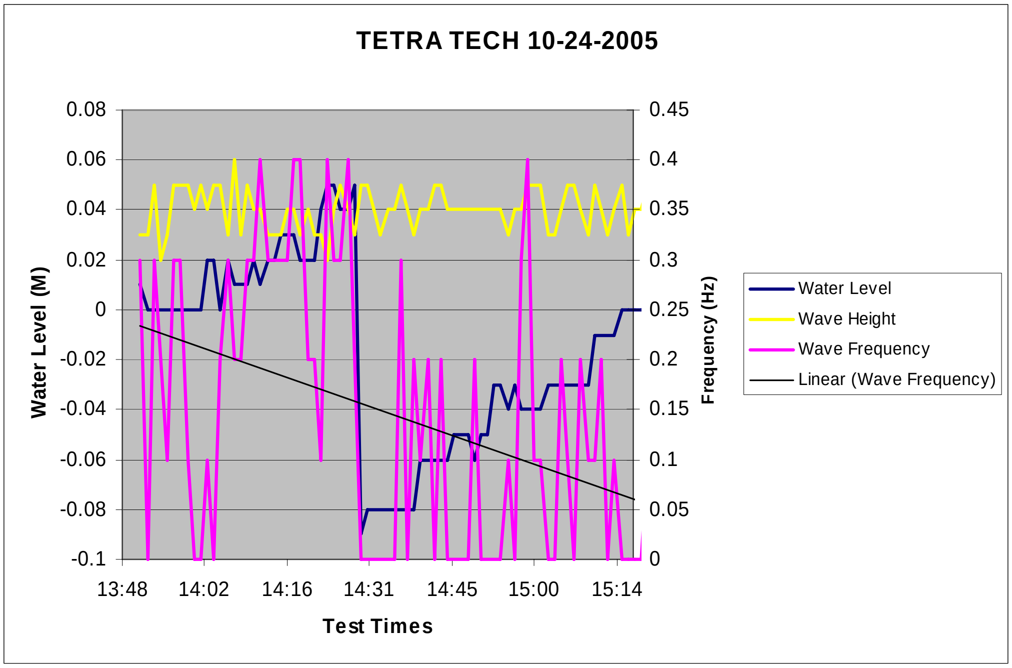

Figure A-1. Water conditions on 24 October 2005.

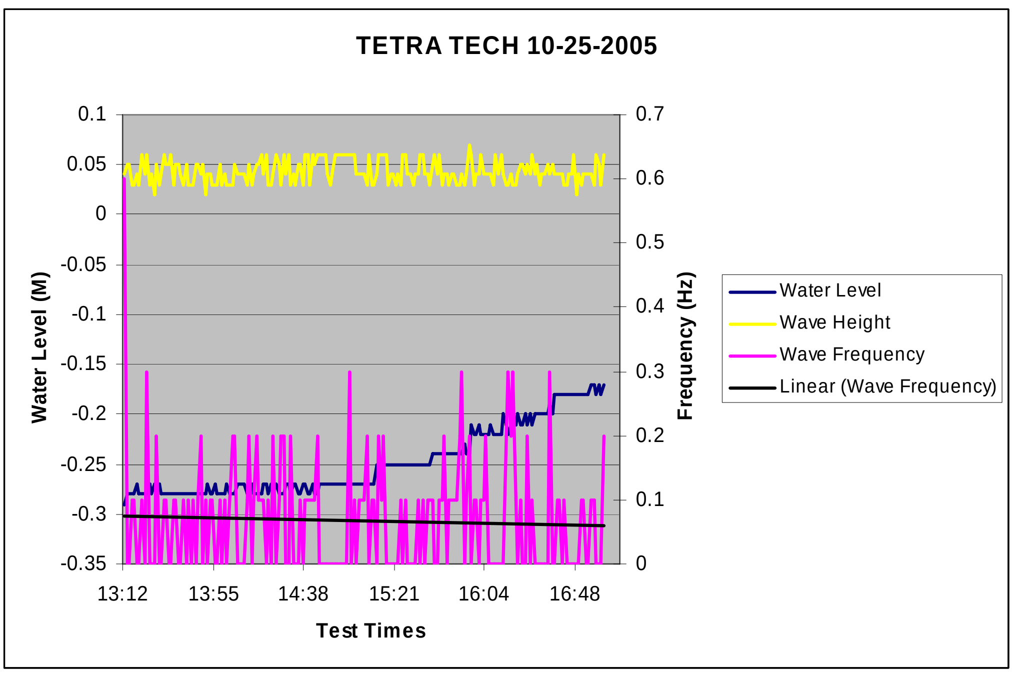

Figure A-2. Water conditions on 25 October 2005.

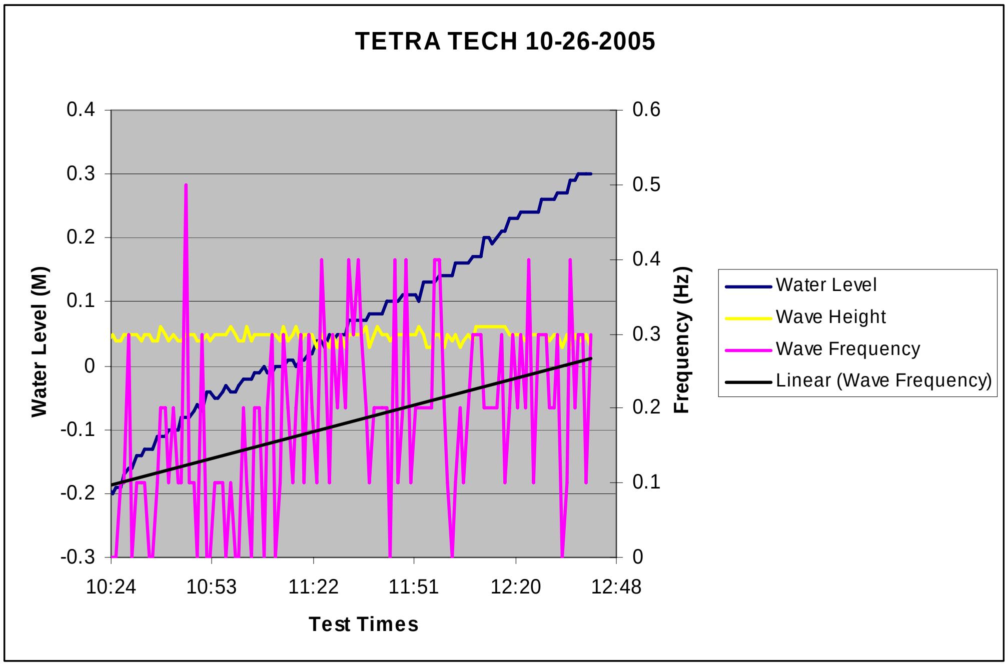

Figure A-3. Water conditions on 26 October 2005.

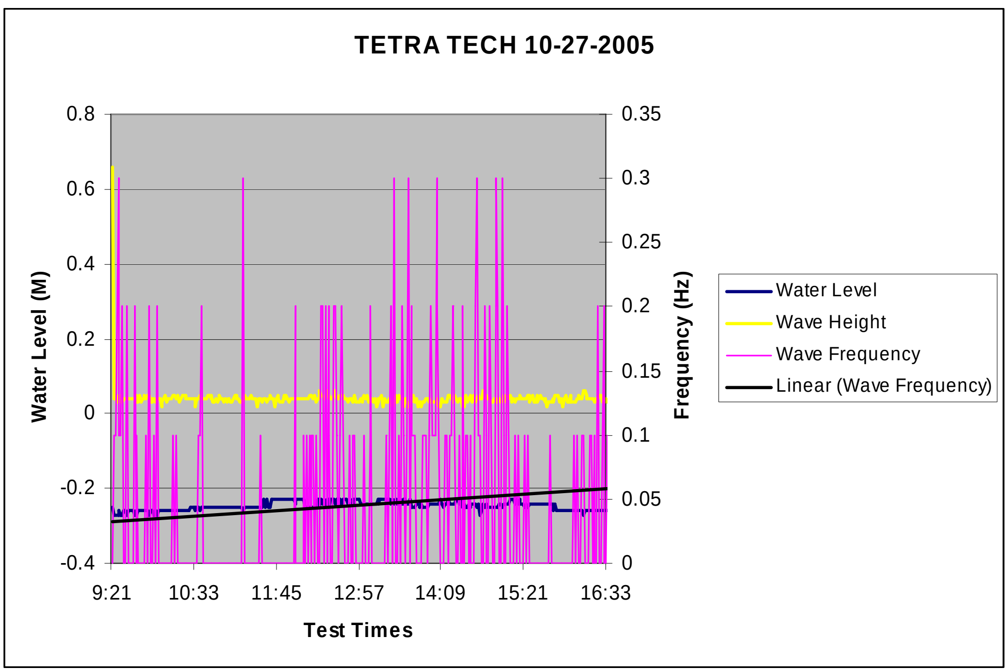

Figure A-4. Water conditions on 27 October 2005.

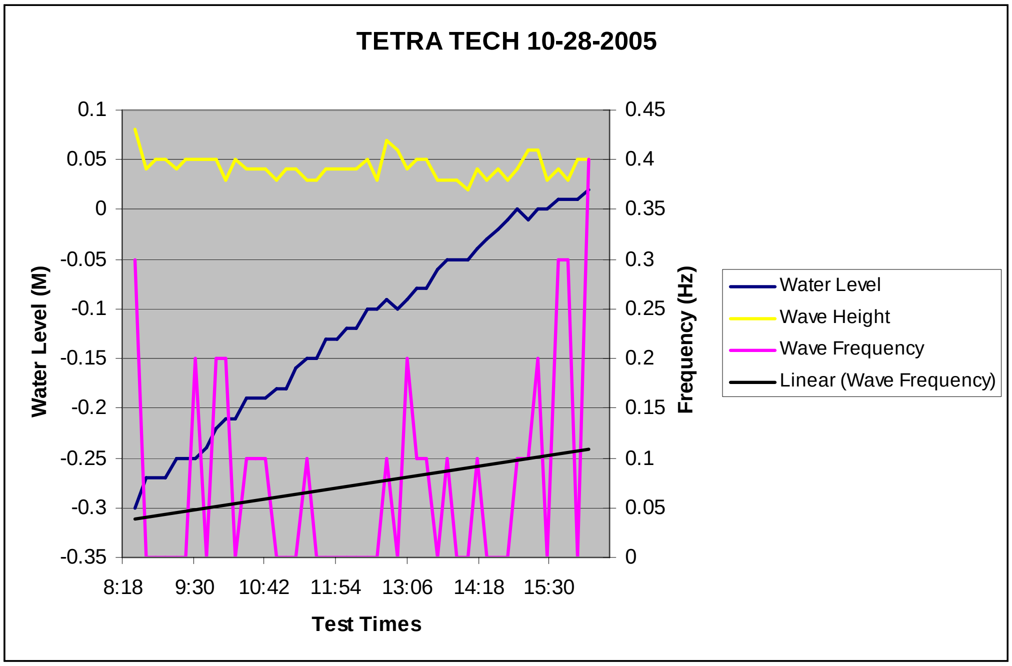

Figure A-5. Water conditions on 28 October 2005.

Water conditions for 29 October 2005 were lost due to an instrumentation malfunction.

Company: Tetra Tech EC, Inc | *Personnel: Tim Deignan, Adam Maiers, Brian Corbett*

Date: 10/24/2005

Start	Stop	Remarks	Activity	Chargeable
0840	0850	Arrived at site. Safety briefing.	Safety briefing	0
0850	1030	Walked area, began setup, waited on a truck to arrive with additional equipment.	Initial setup	100
1030	1345	Truck arrived, off loaded equipment. Continued setup.	Initial setup	195
1345	1405	Tested the imaging sonar in the pond.	Calibration	20
1405	1625	Returned to the dock, began installing magnetometers.	Initial setup	140
1625	1700	MAGS installed. Out on the water for reading check. Seemed to follow same path taken during the sonar check.	Calibration	35
1700	1800	Packed up for the day. Left site.	Daily close-up	60

Date: 10/25/2005

Start	Stop	Remarks	Activity	Chargeable
0800	0915	Arrived at site - bailed rain out of the boat. Currently raining. One cable needs to be changed based on the QC check. Constructed a makeshift shelter from the rain for the equipment and operators.	Daily setup	75
0915	0945	Setup and QC check on background unit.	Calibration	30
0945	1355	Calibration lane and “quit” area of pond surveyed as part of the QC checks.	Calibration	250
1355	1450	Break.	Lunch	45
1450	1515	The electric motors had a difficult time with today’s wind. Conducted a short test using the facility 6-horsepower outboard motor to see if it produced too much interference for use with this system. The answer is yes.	Downtime equipment	25
1515	1615	Surveyed using the two electric motors.	Survey	60
1615	1700	Cleaned up. Left site.	Daily close- up	45

Company: Tetra Tech EC, Inc		*Personnel: Tim Deignan, Adam Maiers, Brian Corbett*		
Date: 10/26/2005				
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0800	1000	Arrived at site/setup.	Daily setup	120
1000	1050	Taking background readings.	Calibration	50
1050	1145	Surveyed.	Survey	55
1145	1315	Returned to dock. One magnetometer had stopped working. Called GEM Systems technical support - fixed the problem on-site.	Downtime equipment	90
1315	1345	Resumed survey.	Survey	30
1345	1350	Same magnetometer stopped working. Reconfigured to survey with 3 instead of 4 magnetometers.	Downtime equipment	5
1350	1500	Surveyed.	Survey	70
1500	1525	Changed motor batteries. TT will order a 3.3 horsepower gas motor to propel the boat. The two electric motors are insufficient to maneuver the system when there is a wind.	Downtime equipment	25
1525	1550	Surveyed.	Survey	25
1550	1640	Cleaned up. Left site.	Daily close-up	50

Date: 10/27/2005				
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0700	0745	Off site this morning; additional maintenance was performed on potassium magnetometer.	Downtime equipment	45
0810	1030	Arrived at site. Reconfigured for a gas motor.	Daily setup	140
1030	1215	Surveyed.	Survey	105
1215	1315	Refueled. Other problems - networking.	Downtime equipment	60
1315	1445	Surveyed.	Survey	30
1445	1455	Refueled.	Downtime equipment	10
1455	1615	Surveyed.	Survey	20
1615	1700	Cleaned up. Left site.	Daily close-up	45

Company: Tetra Tech EC, Inc			*Personnel: Tim Deignan, Adam Maiers, Brian Corbett*	
Date: 10/28/2005				
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0810	0920	Arrived on site, began setup.	Daily setup	70
0920	1000	Experimentation: relocated GPS head above magnetometers to see what the signal may look like at different heights. Mounted GPS head.	Calibration	40
1000	1015	"Walk-away"check.	Calibration	15
1015	1215	Surveyed.	Survey	120
1215	1220	Refueled.	Downtime equipment	5
1220	1320	Surveyed.	Survey	60
1320	1330	Taking measurements of GPS/MAG configuration for data processing.	Calibration	10
1330	1530	Surveyed. Filled in gaps caused by wind during the first days of testing.	Survey	120
1530	1535	Refueled.	Downtime equipment	5
1535	1610	Surveyed.	Survey	35
1610	1650	Problem with sensor No. 1 will continue working on it overnight.	Downtime equipment	40
1650	1730	Cleaned up. Left site.	Daily close-up	40

Date: 10/29/2005				
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0750	0915	Arrived on site, began setup. Reconnected MAG that was worked on last night. Can survey without it using 3 sensors.	Daily setup	85
0915	0945	Sensor checked.	Calibration	30
0945	1115	Surveyed.	Survey	90
1115	1125	Refueled.	Downtime equipment	10
1125	1345	Surveyed.	Survey	140
1345	1355	Refueled.	Downtime equipment	10
1355	1415	Surveyed.	Survey	20
1415	1430	Reviewed data.	Calibration	15
1430	1530	Surveyed. Concentrated on shallow side of the littoral zone. The 2 sonar units were put in the boat.	Survey	60
1530	1650	Demobilization. Left site.	Demobilization	80

APPENDIX C. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Munitions and Explosives Of Concern (MEC): Specific categories of military munitions that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g. TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the projected length of the ordnance onto the ground plane plus 1 meter.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selects the threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open water only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can, therefore, be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can, therefore, be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

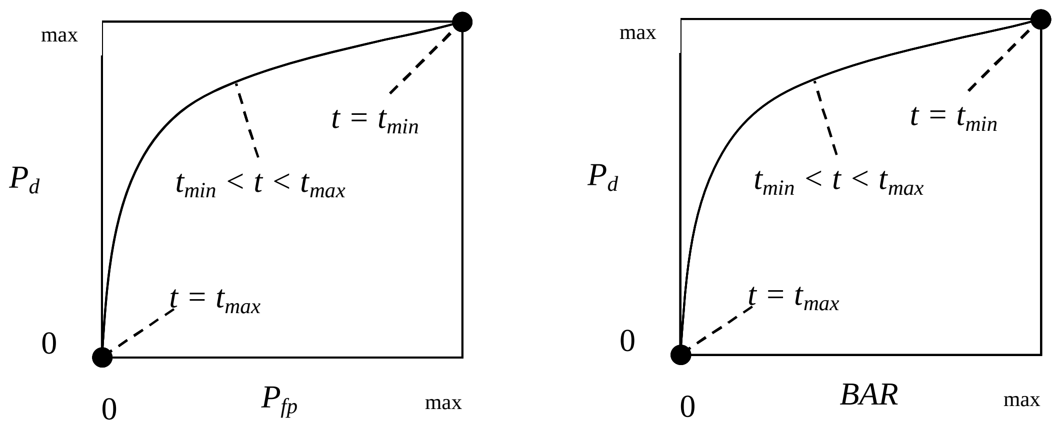

Figure A-1. ROC curves for open-site testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an Open Water scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the Open Water ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$: measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage t_{min}) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$: measures (at a threshold of interest) the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage t_{min}). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$
Open water: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 4, pages 144 through 151).

A one-sided 2 x 2 contingency table is used in the Shallow Water Site Program to compare each area (Open Water, Littoral, Deep Water) to the Blind Grid since each area introduces a water feature that makes it potentially more difficult to survey than the Blind Grid. The one-sided 2 x 2 contingency table is used to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging feature introduced. A two-sided 2 x 2 contingency table is used to compare performance between any two of the test sites other than the Blind Grid, to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly different between those two test sites.

The test statistic of the 2 x 2 contingency table is the Chi-square distribution with one degree of freedom. For the one-sided test, a significance level of 0.05 is chosen which sets a critical decision limit of 3.84 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's Exact Test is used and the critical decision limit is the chosen significance level, which is 0.05 for one-sided tests and 0.10 for two-sided tests. With Fischer's test, if the test statistic (p-value) is less than the critical value, then the null hypothesis of similar performance is rejected in favor of the alternative hypothesis: significantly greater than for the one-sided case or significantly different for the two-sided case.

Shallow-water UXO Detection Test Site examples, where blind grid results are compared to those from the open water and littoral sites and the non-grid sites (open water and littoral) are compared to each other as follows. It should be noted that a significant result does not prove a cause and effect relationship exists between the change in survey area and sensor performance; however, it does serve as a tool to indicate that one data set reflects relatively degraded system performance of a large enough scale than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open water	Littoral
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open water. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic (p-value) of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open water relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 out of 10 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used in the Chi-square Contingency Test to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 3.84, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 100 out of 100 and 20 out of 33 are used to calculate a test statistic (< 0.000) that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.61) is considered to be significantly less at the 0.05 level of significance.

P_d^{disc}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 and 8 out of 33 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used to calculate a test statistic of 32.01. Since the test statistic is greater than the critical value of 3.84, the smaller discrimination stage detection rate (0.24) is considered to be significantly less at the 0.05 level of significance.

P_d^{res}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.10 level of significance.

P_d^{disc}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the two discrimination stage detection rates are considered to be significantly different at the 0.10 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and change in performance, it does indicate that the ability of Demonstrator X to correctly discriminate seems to have been degraded by features of the littoral area relative to results from the open water using the same system.

APPENDIX D. REFERENCES

1. Environmental Quality Technology - Operational Requirements Document (EQT-ORD) for: A(1.6.a): UXO Screening, Detection and Discrimination.

2. Technical Management Plan, Unexploded Ordnance Detection and Discrimination Demonstration for the APG Standardized UXO Technology Shallow Water Demonstration Site. Submitted in response to the BAA W91ZLK-04-R-0001 by Tetra Tech EC, Inc., 9 August 2005.

3. *Email*: 6 July 2006, sent from Mr. Timothy Deigan (timothy.deigan@tteci.com) to Mr. Gary Rowe (gary.rowe@atc.army.mil) regarding SWDS Scoring.

4. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX E. ABBREVIATIONS

APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
BAA	=	Broad Agency Announcement
BAR	=	background alarm rate
DGPS	=	Differential Global Positioning System
DMM	=	discarded military munitions
EQT	=	Army Environmental Quality Technology Program
EQT-ORD	=	Environmental Quality Technology - Operational Requirements Document
ERDC	=	U.S. Army Corps of Engineers Engineering, Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
GPS	=	Global Positioning System
HFIS	=	High Frequency Imaging Sonar
LED	=	light-emitting diode
MEC	=	munitions and explosives of concern
MEDTC	=	Military Environmental Technology Demonstration Center
MFSBP	=	Multiple Frequency Sub-Bottom Profiler
MGS	=	magnetic gradiometer system
NMEA	=	National Marine Electronics Association
P_{ba}	=	probability of background alarm rate
P_d	=	probability of detection
P_d^{disc}	=	probability of detection, discrimination stage
P_d^{res}	=	probability of detection, response stage
P_{fp}	=	probability of false positive
P_{fp}^{disc}	=	probability of false positive, discrimination stage
P_{fp}^{res}	=	probability of false positive, response stage
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver operating characteristics
RTK	=	real-time kinematic
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Center
UXO	=	unexploded ordnance

APPENDIX F. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-016

Addressee	No. of Copies
Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT (Ms. Bonnie Packer) Aberdeen Proving Ground, MD 21010-5401	1
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E (Mr. Gary Rowe)	1
(Library)	1
CSTE-DTC-AT-CS-R Aberdeen Proving Ground, MD 21005-5055	1
Tetra Tech EC, Inc ATTN: (Mr. Tim Deignan) 143 Union Blvd, Suite 1010 Lakewood, CO 80228	1
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary, "hard-copy" distribution is controlled by Commander, U.S. Army Environmental Center, ATTN: SFIM-AEC-ATT.